

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-028817

**(43)Date of publication of application : 31.01.1995**

(51)Int.Cl.

G06F 17/27

G06F 17/21

(21)Application number : 05-175108

(71)Applicant : HITACHI LTD

(22)Date of filing : 15.07.1993

(72)Inventor : UEHARA TETSUZO

(54) STRUCTURE CONVERTING METHOD OF DOCUMENT INFORMATION

**(57)Abstract:**

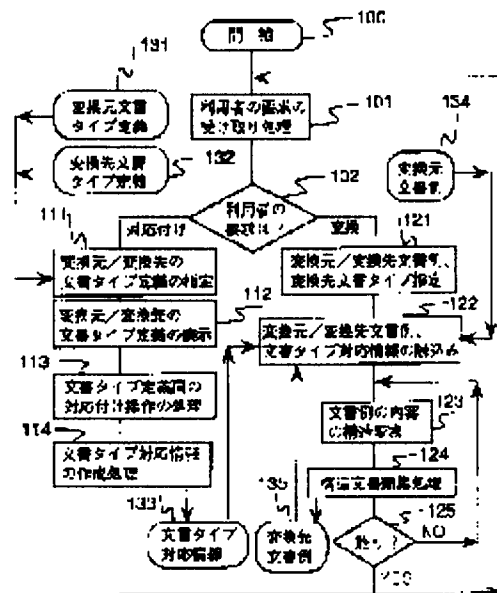
**PURPOSE:** To provide a structure converting method of document information which is capable of converting a document example conforming to a document type A into the document example conforming to a document type B by designating the correspondence relation between two kinds of A and B of document type definitions which define the structure of the document, preparing correspondence information and utilizing this.

CONSTITUTION: A conversion origin document type definition 131

and a conversion destination document type definition 132 including

the description of the document type definition sentence of various kinds of documents are held,

document type correspondence information 133 is prepared by the corresponding operation processing



113 between document type definitions for a document type correspondence designated example designating the correspondence relation between these definitions and the preparation processing 114 of document type correspondence information, this information and a conversion origin document example 134 are made input information, a conversion destination document example 135 is prepared by the structure conversion processing 123 of the contents of the document example and the document example is imparted to a structure document edition processing 124.

(51) Int.Cl.<sup>6</sup>

識別記号

庁内整理番号

F I

技術表示箇所

G 0 6 F 17/27

17/21

7315-5L

G 0 6 F 15/ 20

5 5 0 E

7315-5L

5 8 0 L

審査請求 未請求 請求項の数 8 O L (全 10 頁)

(21) 出願番号 特願平5-175108

(22) 出願日 平成5年(1993)7月15日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 上原 徹三

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

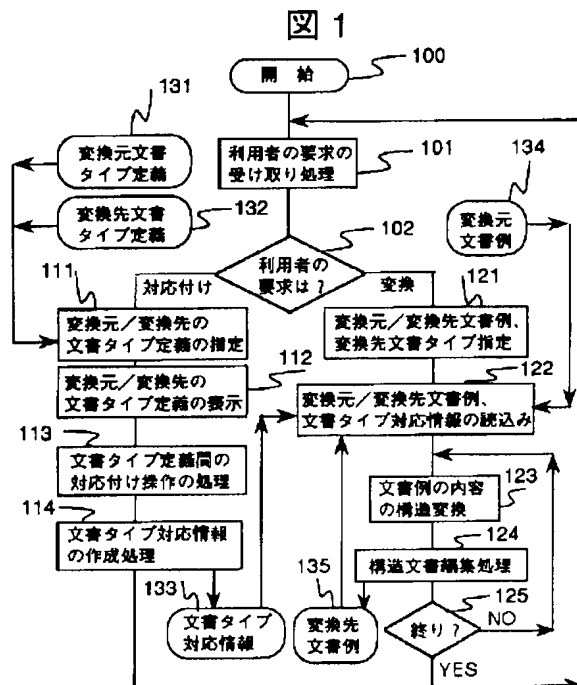
(74) 代理人 弁理士 小川 勝男

(54) 【発明の名称】 文書情報の構造変換方法

(57) 【要約】

【目的】 文書の構造を定義したA、B二種の文書タイプ定義間の対応関係を指定して対応情報を作成し、これを利用して、文書タイプAに従う文書例を文書タイプBに従う文書例に変換できるような文書情報の構造変換方法を提供する。

【構成】 各種の文書の文書タイプ定義文の記述を含む変換元文書タイプ定義131および変換先文書タイプ定義132を保持し、これらの間の対応関係を指定する文書タイプ対応付け指定例に対する文書タイプ定義間の対応付け操作の処理113と文書タイプ対応情報の作成処理114によって、文書タイプ対応情報133を作成し、これと変換元文書例134を入力情報として、文書例の内容の構造変換処理123により変換先文書例135を作成すると共に、それを構造文書編集処理124で与える。



## 【特許請求の範囲】

【請求項1】複数の文書に共通の文書の形式を、その構成単位とその出現順序と出現個数によってを記述し、前記構成単位の形式を同様に記述することを適宜繰返すことによって、文書の形式を定義したデータである文書タイプ定義として、第1の文書タイプ定義と、第2の文書タイプ定義との2種を与えられて、これらの文書タイプ定義間の利用者による対応付け操作によって、前記第1の文書タイプ定義に従う文書例を前記第2の文書タイプ定義に従う文書例に変換する際に利用する文書タイプ対応情報を作成することを特徴とする文書情報の構造変換方法。

【請求項2】請求項1に記載の利用者による文書タイプ定義間の対応付け操作として、前記第1の文書タイプ定義の構成単位と前記第2の文書タイプ定義の構成単位との間の1:1の対応関係を与える文書情報の構造変換方法。

【請求項3】請求項1に記載の利用者による文書タイプ定義間の対応付け操作として、前記第1の文書タイプ定義の含む構成単位の一部の構成単位と前記第2の文書タイプ定義の構成単位の含む構成単位の全体または一部の構成単位との間の1:1の対応関係を与える文書情報の構造変換方法。

【請求項4】請求項1に記載の利用者による文書タイプ定義間の対応付け操作として、前記第1の文書タイプ定義の含む構成単位の出現順序を前記第2の文書タイプ定義の構成単位の対応する構成単位の出現順序とが異なるような対応関係を与える文書情報の構造変換方法。

【請求項5】請求項1に記載の利用者による文書タイプ定義間の対応付け操作として、前記第1の文書タイプ定義の含む構成単位の系列と前記第2の文書タイプ定義の構成単位との間の対応関係を与える文書情報の構造変換方法。

【請求項6】請求項1に記載の利用者による文書タイプ定義間の対応付け操作において、前記第1の文書タイプ定義の一部の構成単位の定義として、出現個数がある一定の複数値であるか、あるいは、任意の複数个数でよいという記述がなされている場合、前記第1の文書タイプ定義の含む構成単位あるいは構成単位系列を、前記第2の文書タイプ定義の前記複数个数指定の構成単の一箇分の出現例と解釈されるような対応関係を与える文書情報の構造変換方法。

【請求項7】請求項1、2、3、4、5または6において、所属する文書タイプの指定された第1の文書例を与えられて、これを別の文書タイプ定義に従う第2の文書例に変換するに際して、前記第1の文書例の文書タイプ定義を前記第1の文書タイプ定義とし、前記第2の文書例の文書タイプ定義を前記第2の文書タイプ定義として作成された、前記文書タイプ対応情報を参照し、この対応関係に応じて、前記第1の文書例の構成単位を、場合

により無視し、場合により順序を変え、場合により変換先文書定義の構成単位として扱うことによって、前記第1の文書タイプ定義に従う文書例を前記第2の文書タイプ定義に従う文書例に変換する文書情報の構造変換方法。

【請求項8】請求項7において、ある文書タイプ定義に従う文書例全体の内容、あるいは、その文書を構成する構成単位に対する文書例の内容が与えられたとき、それら与えられた内容を指定の文書タイプ定義に含まれる構成単位と対応付けることによって、文書タイプ定義に従っていることを確認しつつ文書例を編集する構造文書編集機能を保持し、文書情報の構造変換において、文書例の内容を、場合により無視し、場合により順序を変え、場合により前記第2の文書定義の構成単位に取り替えることによって得られる文書内容を、前記第2の文書タイプ定義に従う文書例の内容であるとして、対応する構成単位の識別と共に前記構造文書編集機能に与えることによって、変換先の文書例の編集を行う文書情報の構造変換方法。

## 【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は構造の定義された文書情報の間の変換方法に関する。

【0002】

【従来の技術】従来の技術(1)として、アイエスオー(ISO:International Standard Organization)による国際規格であるISO8879:インフォメーションプロセッシング……テキスト アンド オフィス システムズ……スタンダードジェネライズド マークアップ ランゲージ(Information processing…Text and office systems…Standard Generalized Markup Language:SGML)(ISO,1986)(以下では、SGMLと略す)が挙げられる。これによると、文書の論理構造を木構造と等価な階層構造として表現する方法が示されている。また、複数の文書間に共通の構造と構成単位の取り得る属性を文書タイプ定義ドキュメント タイプ デフィニション(DTD:Document Type Definition)として表現し、個々の文書はその所属するタイプを指定するとともに、その文書タイプ定義の許容する構成と属性を持ち、また、その文書タイプ定義の許容する属性に対して属性値を設定することによって表現される。

【0003】従来の技術(2)は、文書タイプ定義DTDに従って、文書内容を入力し、編集することを支援する構造文書編集機能の実現例がある。その一例として、ソフトクウォード(SoftQuad)社のプログラム製品であるオーザ エディタ(AuthorEditor)が挙げられる。これは文書タイプ定義を指定して文書の入力編集を行い、指定した文書タイプ定義の構成単位ごとに文書例の内容を入力できる。文書タイプ定義の規定した構成に違反するような、文書内容の作成や修正操作は拒否される。こ

れにより、文書タイプ定義に従った文書内容の作成を効率的に行うことが可能である。

【0004】従来の技術(3)は、ラテック(LaTeX)などのように文書情報を割り付ける方法を指定する割り付けスタイル情報を用意し、割り付け時に異なる割り付けスタイル情報を用いると、同一の文書の割り付け形式が変更できるという機能がある。これにより、文書本体の変更なく、文書の出力形式を変更できる。

【0005】

【発明が解決しようとする課題】従来の技術(1)は、文書の論理構造の表現方法を示しており、また、文書のクラスの定義、即ち、文書タイプ定義により、それに従う個々の文書例の形式を表現する考え方を示している。しかし、文書タイプ間の関係については規定していない。

【0006】従来の技術(2)は、文書タイプ定義に従った文書例の作成の効率化と、その可視化方法や、割り付け方法との対応付けの機能を実現しているが、異なる文書タイプ定義に従う文書間の変換に関する機能は実現していない。

【0007】従来の技術(3)は、同一の論理的内容の文書に対して、割り付けスタイル情報の変更によって、異なる割り付け形式の出力を用容易に与えられる。しかし、論理的構造の異なる文書間の構造の変更を行うものではない。

【0008】今後、前記のようなSGML文書が普及することが予想されるが、その場合、文書の複数種類の(論理)構造定義である文書タイプ定義とこの文書タイプ定義ごとにまた多数の文書例が作成されるようになる。すなわち、文書タイプ定義DTD(a)の各々に文書例Doc(a-j)が存在する。これらの文書タイプDTD(a)とDTD(b)とは、全く異なる文書の文書タイプ定義の場合もあるが、関係する文書の文書タイプ定義の場合も十分あり得る。例えば、次のような場合がある。

【0009】A) DTD(a)とDTD(b)とは構造上は全く等価で、構造単位の名前のみが異なる。

B) DTD(a)はDTD(b)の部分の並べ変え。

C) DTD(a)はそれに特有の部分とDTD(b)の複数文書例の全体とから構成される。

D) DTD(a)はそれに特有の部分とDTD(b)の複数文書例の部分とから構成される。

【0010】さて、これらの文書タイプに従う文書例が、少なからず作成されると、関連する文書タイプ相互の間での変換の要求が発生する。例えば、ある学会用に作成した論文を別の学会用の論文に変更したい場合

(A)の例参照)、ある論文誌の論文のタイトルと著者名と要旨のみを抜き出した要約を作成したい場合(B)の例参照)、ある学会の論文記述要領に従ったいくつかの論文を集めるとともに、先頭および最後に、別の形式

の文章を付加することによって、1冊の論文誌を構成したい場合(C)の例参照)、ある論文誌に記載された複数の論文のタイトルと著者名と要旨のみを抜き出した要約集を作成したい場合(D)の例参照)などの要求である。これらの要求は、一括して、文書例の間の文書情報の構造変換の要求であるということが出来る。

【0011】本発明の目的は、変換元文書例の従う文書タイプ定義および変換先文書例の従うべき文書タイプ定義を利用して、文書情報の構造変換を行い、変換元文書例から変換先文書例を作成することにある。

【0012】本発明の第1の目的は、変換元文書タイプ定義および変換先文書タイプ定義の間に利用者が対応付けを指定する方法を提供することである。より詳しくは、上記のような種々の場合に対処すべく、変換元文書タイプ定義の構成単位と変換先文書タイプ定義の構成単位との間の1:1の対応関係、変換元文書タイプ定義の含む構成単位の一部の構成単位と変換先文書タイプ定義の構成単位の含む構成単位の全体または一部の構成単位との間の1:1の対応関係、変換元文書タイプ定義の含む構成単位の出現順序と変換先文書タイプ定義の構成単位の対応する構成単位の出現順序とが異なるような対応関係、変換元文書タイプ定義の含む構成単位の系列と変換先文書タイプ定義の構成単位との間の対応関係を与えることを可能とすることを目的とする。また、変換先文書タイプ定義の一部の構成単位の定義として複数個数出現の記述がなされている場合、変換元文書タイプ定義の含む構成単位あるいは構成単位系列を、変換先文書タイプ定義の前記複数個数指定の構成単位の一つ分の出現例と解釈されるような対応関係を与え得ることを目的とする。

【0013】本発明の第2の目的は、実際に与えられた変換元文書例を、このように利用者が指定した対応付けに従って変換し、変換先文書例を作成することである。更に、このように変換した文書内容を公知の構造文書編集機能に渡すインターフェースを提供することである。

【0014】

【課題を解決するための手段】上記の第1の目的に対して、本発明は、変換元の文書例の従う文書タイプ定義である変換元文書タイプ定義と変換先の文書例の従う文書タイプ定義である変換先文書タイプ定義とを用意するとともに、利用者による文書タイプ定義間の対応付け操作手段を提供し、また、その指定結果を示す文書タイプ対応情報の作成手段を提供する。

【0015】上記の第2の目的に対して、本発明は、変換元文書例と変換先文書タイプを指定し、該当する文書タイプ対応情報を参照し、変換先文書例を作成する手段を提供する。また、作成した変換先文書の構成単位の内容を構造文書編集機能に渡す手段を提供する。

【0016】

【作用】上記の第1の目的を実現するために、本発明で

は、変換元の文書例の従う文書タイプ定義である変換元文書タイプ定義と変換先の文書例の従う文書タイプ定義である変換先文書タイプ定義とを用意すると共に、利用者による文書タイプ定義間の対応付け操作を用いて、適当な変換元文書タイプ定義の構成単位あるいはその系列と、該当する変換元文書タイプ定義の構成単位との間の対応関係を指定することにより、両文書タイプの構造上の関係を与えることができる。この対応関係の指定は、先に記した、A) からD) の場合に相当する対応関係を含む指定が可能である。なお、これらの指定結果に従って、文書タイプ対応情報を作成し、変換元文書タイプの識別と定義変換先文書タイプ定義の識別とのペアによって取り出される形で保存し、下記の変換処理で利用できるようにする。

【0017】第2の目的を実現するために、本発明では、変換元文書例と変換先文書タイプを指定されて、変換元文書タイプと変換先文書タイプの識別により、該当する文書タイプ対応情報を取り出して参照し、これと変換元文書例の内容とを突き合わせつつ変換先文書例を作成することができる。また、この変換先文書例の構成単位を作成することにより、その内容を構造文書編集機能に渡すことによって、変換先文書例を変換しながら編集、表示等の公知の構造文書編集機能の文書処理機能を利用すること、および、既存の変換先文書例に変換元文書例からの内容を変換して追加することを可能とする。

【0018】

【実施例】図2は本発明の文書情報の構造変換方法を提供する文書システムの一構成例である。図2で、処理装置20は、本システムを実現するプログラムを実行する処理装置であり、キーボード21は、本システムの利用者が文書名を指定したり、システムからの問い合わせに応えたりするための文字入力手段であり、ディスプレイ装置22は、文書例や文書タイプ定義を表示する出力装置であり、指示入力装置23は、ディスプレイ装置22に表示された部分を指示することにより、利用者からの指定を与える入力装置であり、外部記憶装置25は、文書例や文書タイプ定義および文書タイプ対応情報などの文書データの保存、本システムのプログラムの保存手段であり、主記憶装置24は、本システムのプログラムおよび文書データを必要に応じて外部記憶装置26からここに読み込み、プログラムによって参照したり、データを生成・変更したりするためのいわゆる主記憶装置である。

【0019】本システムを起動すると、本システムを実現するためのプログラムが、外部記憶装置25から主記憶装置24に読み込まれ、このプログラムに制御が渡される。

【0020】図1によりこのプログラムの動作を説明する。このプログラムは、本発明の文書情報の構造変換方法を与えるものである。プログラムは、符号100の部

分から開始されると、まず、処理101で、ディスプレイ装置22へ表示によって、利用者に対してその行いたいことを問いかけ、その要求の指定を持つ。利用者は、キーボード21あるいは指示入力装置23を用いてその要求を指定する。ここでの利用者の要求とは、文書タイプ定義間の対応付けの場合と、文書内容の構造変換の場合との二種がある。処理102はこれらを判別し、前者であれば処理111に進み、後者であれば処理121に進む。処理111から始まる文書タイプ定義間の対応付け処理とは、入力情報である変換元文書タイプ定義131と変換先文書タイプ定義132とに関して、利用者に対して両文書タイプ定義間の対応付け操作を行わせ、利用者による操作内容によって、文書タイプ対応情報133を作成して外部記憶装置25に保存する。

【0021】ここで、以下の処理の記述の前提となる文書タイプ定義について、まず、例を挙げて説明する。図3と図4は、文書情報の文書タイプ定義の例を与える図である。このうち、図3は、文書の一種である論文に関する文書タイプ定義の例である。論文にも色々な形式があり、唯一つの文書タイプ定義で定義できない。しかし、論文としてまとめられる以上、共通の性質もある。図3は、二つの論文の種類である論文Aと論文Bに関する文書タイプ定義の例を示したものである。図3(a)が論文Aに関する文書タイプ定義、図3(b)が論文Bに関する文書タイプ定義を示す。なお、文書タイプ定義の全体は規模が大きく、ここではその一部を示した。

【0022】図3の(a)の論文Aに関する文書タイプ定義31について説明する。文書タイプ定義は、いくつかの定義文からなる。文書タイプ定義31で、A. 1)、A. 2)、A. 3)などと記した各々が定義文の例である。〈〉で囲んで示したのがこの文書の構成単位であると考ええる。即ち、〈タイトル〉とは、タイトルという名称の構成単位を指す。一々、タイトルという名称の構成単位と言うのは面倒なので、説明でも〈タイトル〉という記号を使うことにする。定義文A. 1)は、論文Aという文書(その最上位の構成単位が〈論文A〉である)の形式が、次の種類の構成単位がその順に出現したものであることを示す。即ち〈論文A〉とは、〈タイトル〉、〈著者〉、〈要旨〉、〈始めに〉、〈本文〉、〈終りに〉、および、〈参考文献〉の系列である。

【0023】これは、普通の言葉で言えば、論文Aの形式は、まず、タイトルがあり、次に著者が書いてあり、その次に、要旨が来て、始めにという項目の後に、本文を記し、終りにという項目を置いて、最後に、参考文献を記す、というような書き方の規定に相当する。定義文A. 1)の記述は、このような規定をより簡明に示す方法である。なお、この記述法は、プログラム言語の文法の記述法として従来採用されているものと、基本的には同じものである。

【0024】定義文A. 2)は、定義文A. 1)で用いられた〈本文〉を定義している。即ち、〈本文〉とは、〈章〉\*である。ここで、\*はその直前の記述の1個以上の出現を表わす。つまり、〈本文〉は、〈章〉が1個以上繰り返し出現したものである。文書タイプ定義は、このように、各構成単位を定義し、その定義に用いた構成単位を更に定義するという風に行われる。しかし、定義文A. 3)以降へと続く論文Aの定義の詳細は、本発明の理解の上で必要でないので図3の(a)では省略した。

【0025】次に、図3の(b)の論文Bに関する文書タイプ定義32を説明する。定義文B. 1)は、〈論文B〉が、〈タイトル〉、〈著者〉、〈序論〉、〈本文〉、〈結論〉、〈謝辞〉?、および、〈参考文献〉の系列であることを示す。ここで〈謝辞〉?とは、〈謝辞〉が有っても無くてもよいことを示す。定義文B. 2)は、論文Aの場合と同様、〈本文〉とは〈章〉\*である、ということを示している。定義文B. 3)以降へと続く論文Bの定義の詳細は省略した。

【0026】もちろん、この定義からは判らないが、ここでは、論文Aの〈始めに〉の定義内容と論文Bの〈序論〉の定義内容とは同じであるとする。〈終りに〉と〈結論〉についても同様である。ここで、論文Aと論文Bの形式を比較しておくと、論文Aには〈謝辞〉がなく、論文Bには〈要旨〉がない。また、論文Aにある〈始めに〉、〈終りに〉は、それぞれ、論文Bにある〈序論〉、〈結論〉に対応するのであるが、両論文でそれらの構成単位の識別名が違っている例である。

【0027】これらの構成単位の識別名の違いを除けば、二つの文書タイプ論文Aと論文Bの形式の相違は、以下の点だけである。論文Aには定義した位置に〈要旨〉があり、一方、〈謝辞〉という構成単位はなく(多分、誰かに感謝したければ、〈終りに〉の末尾にでも記すのである)、論文Bには定義した位置に〈謝辞〉を置くことができ(但し、省略してもよい)、一方、〈要旨〉がない。即ち、二つの文書タイプ論文Aと論文Bとは、極めて類似の文書タイプを定義していることになる。そして、SGMLの普及などによって、文書タイプの定義が広く用いられるようになれば、このようなことは、実際に十分起こり得ることである。可能ならば、文書タイプの統一、標準化が望ましい場合もあるが、既存の文書は従来その形式を持っているのであり、現実には、本来異なる文書タイプの定義を統一するよりも、むしろ、類似の文書タイプ定義の間の関係を取り扱う本発明のような機能が要求されると考えられる。

【0028】例えば〈始めに〉と〈序論〉の例のように、内容上は等価でありながら構成単位の識別名が異なるという事態も、論文Aと論文Bの文書タイプ定義が独立になされ得る(あるいは、既に過去に独立になされた)ことだけからでも、十分起こり得ることであること

がわかる。

【0029】また、文書タイプを新規に定義するにしても、内容上は等価な構成単位の識別名を異なるものにしたという要求はあり得る。例えば、日本語の文書では日本語の構成単位名を、英語の文書では英語の構成単位名を用いたいであろう。子供用の文書と大人用の文書では構成単位名も、それにふさわしい識別名にしたいであろう。先に挙げた構造文書機能を用いて文書を作成する時、構造文書機能の側から、作成対象文書の文書タイプ定義に従った構成単位の識別名の系列が表示され、文書作成者は、ある構成単位の内容を入力したり、表示させたり、文字サイズや書体を変えたりできる。この場合にも、自らの作成する文書のタイプにふさわしい構成単位の識別名を用いることが、利用者の文書作成効率にとって望ましい。このように、類似の文書タイプであっても、その文書タイプ定義を完全に統一することは、現実問題として可能でもなく、また、望ましいことでもない。しかし、一方、類似の文書であるからには、共通の扱いをしたいことが起こるのは当然である。これを文書タイプ間の関係付けによって実現する。

【0030】次に図4により、別の文書タイプ定義の例を説明する。図4(a)は論文抄録Cの文書タイプ定義、図4(b)は論文誌Dの文書タイプ定義を示す。図4(a)の論文抄録Cの文書タイプ定義41を見ると、定義文C. 1)で、〈論文抄録C〉とは〈全タイトル〉と〈抄録〉\*との系列であることが示され、この〈抄録〉は定義文C. 2)により〈タイトル〉、〈著者〉、〈要旨〉の系列であることが示されている。定義文C. 3)の〈全タイトル〉の定義以降の記述は省略した。

【0031】この定義から判ることではないが、実は論文抄録Cの〈抄録〉部分は、先に定義した〈論文A〉の構成単位である〈タイトル〉、〈著者〉、〈要旨〉の系列であるものとする。この〈抄録〉に\*が付加されているから、この構成単位は繰り返されるのであり、〈論文A〉の異なる文書例の対応部分がここに取り込まれる。

【0032】〈全タイトル〉はこの〈論文抄録C〉全体のタイトルである。即ち、〈論文抄録C〉は、そのタイトルの後に、〈論文A〉のいくつかの文書例の〈タイトル〉、〈著者〉、〈要旨〉の系列を繰り返して配置したものであり、実際によく見られる論文の抄録誌の形式を示している。

【0033】次に、図4(b)の論文誌Dの文書タイプ定義を簡単に見ておこう。まず、定義文D. 1)は〈論文誌D〉が〈全タイトル〉と〈序文〉と〈論文A〉\*の系列であることを示す。定義文D. 2)は〈論文A〉、定義文D. 3)は〈全タイトル〉、定義文D. 4)は〈序文〉の定義をそれぞれ与えるが、それらの定義内容と以降の定義内容については省略した。

【0034】実は、この〈論文誌D〉は、〈論文A〉の

文書例をいくつか配置したものの前に、〈全タイトル〉と〈序文〉を付加したもので、よくある学会誌の形式を単純化したものである。注意すべきは、図4の(a)で文書タイプ〈論文抄録C〉の構成単位が別の文書タイプ〈論文A〉の構成単位を含み、図4の(b)で文書タイプ〈論文抄録D〉の構成単位が、別の文書タイプ〈論文A〉の構成単位全体を含むことであり、これらは、先に図3の説明の後に言及した本発明の対象とする文書タイプ間の関係の別の例である。

【0035】図1に戻り、処理111から始まる文書タイプ定義間の対応付け処理について説明する。処理111は、利用者がこれから対応付け指定を行いたい変換元／変換先の文書タイプ定義の識別名をキーボード21により指定し、プログラムがそれらを受け取る処理である。この後、処理112に進み、ディスプレイ装置22の画面に、上で指定された変換元／変換先の文書タイプ定義を表示する。

【0036】この画面表示の例を、図5によって説明する。図5の(a)は、処理111で、変換元文書タイプとして論文A、変換先文書タイプとして論文Bを利用者が指定したのに対して、処理112によりディスプレイ装置22に表示される文書タイプ対応付け操作画面(対応付け画面と呼ぶ)の例である。この対応付け画面の内容510は、外部記憶装置25から利用者の指定した識別名によって取り出した変換元文書タイプ定義131の内容520および変換先文書タイプ定義132の内容530とから成る。この表示内容は、それぞれ、図3の(a)および(b)に示した論文Aと論文Bの文書タイプ定義の第1の定義文の内容を示したものである。但し、それぞれの構成単位の前に〔 〕を表示すると共に、各構成単位を行を変えて表示している。図5の(b)は、処理111で、変換元文書タイプとして論文A、変換先文書タイプとして論文抄録Cが指定されたときに、処理112により表示される対応付け画面の例である。この内容550のうち、左側の内容560が変換元文書タイプ定義、右側の内容570が変換先文書タイプ定義を表わすことなど、図5の(a)と同様である。

【0037】図1の処理フロー処理112で、上のような対応付け画面が表示された後、処理は、次の処理113に移る。処理113は、この対応付け画面の入力領域である構成単位の前〔 〕の中に、利用者が文書タイプ定義間の対応付け指定を行い、プログラムがそれを受け取るものであり、文書タイプ定義間の対応付け操作の処理という。

【0038】図6により、利用者による対応付け指定の例と、その意味する所を説明する。図6の(a)は、図5の(a)の対応付け画面に対して、利用者が指定した対応付け指定の内容610を示す。この内容は、図5の(a)の左右の文書タイプ定義の各構成要素に対応する〔 〕の中に、利用者がそれらの対応付けを示す記号

(ここでは、a, b, cなどの文字によっている)を入力したものである。この対応付けの方法は、変換元／変換先文書タイプ定義の構成要素に付された記号が同じものを対応させるものである。

【0039】例えば、図6の(a)で、変換元文書である論文Aの構成要素〈タイトル〉と、変換先文書である論文Bの構成要素〈タイトル〉の両者に〔a〕と付されているからこれらは対応する。なお、両者の構成要素の識別名はこの対応関係にとって意味を持たないことに注意する。つまり、識別名が同じだから対応する、識別名が異なるだから対応しない、ということはない。すべては、ここでの利用者の指定による。

【0040】図6の(a)では、〈タイトル〉と同様に、それぞれの文書の構成要素である〈著者〉、〈本文〉、〈参考文献〉の対応が指定されている。更に、変換元文書である論文Aの構成要素〈始めに〉と、変換先文書である論文Bの構成要素〈序論〉には、共に〔c〕と付されているからこれらは対応する。〈終りに〉と〈結論〉も同様である。つまり、先に図3の説明と共に述べた論文Aと論文Bの形式の類似性はこの対応付けの指定内容によってプログラムに対して示される。それらの相違点は、前者が含む〈要旨〉を後者が含まないこと、逆に、後者が含む〈謝辞〉を前者が含まないこと、前者の構成単位の識別名〈始めに〉と〈終りに〉と〈結論〉に対して、後者ではそれぞれ識別名〈序論〉と〈結論〉を用いていることである。図6の(a)の利用者の指定はこの対応関係を示している。また、注意すべきは、このような指定によって、両文書タイプがこのような関係を持つことが初めて明らかになることである。

【0041】図6の(b)は、図5の(b)の対応付け画面に対して、利用者が指定した対応付け指定の内容650を示す。内容650の、変換元文書タイプ定義の部分660の表示内容は(a)と同様である。変換先文書タイプ定義の部分670の表示内容は、図5の(b)の対応付け画面の該当部分と少し異なっている。即ち、〈抄録〉に対する定義文の内容が更に、画面に加えられている。これは、例えば、利用者が図5の(b)の対応付け画面において、〈抄録〉の部分で指示入力装置23によって指示しつつ、その定義情報を出力するように要求したのに対して表示されたものである。図6の(b)は、この後で利用者が指定した対応付け指定の内容650を示したものである。その指定方法は先に述べた通りである。その指定内容は、変換元文書タイプの〈論文A〉の構成単位である〈タイトル〉と〈著者〉と〈要旨〉とが、それぞれ、変換先文書タイプの〈論文抄録C〉の構成単位である〈抄録〉を構成する構成単位である〈タイトル〉と〈著者〉と〈要旨〉とに対応することを示す。

【0042】即ち、先述のように、〈論文抄録C〉は、そのタイトルの後に、〈論文A〉のいくつかの文書例の



〈タイトル〉、〈著者〉、〈要旨〉の系列を繰り返して配置したものである。繰り返して言うが、図6の(b)の利用者の指定により、初めてこの対応関係が明らかになるのである。なお、図には示さなかったが、図4の〈論文誌D〉がそれ独自の構成単位である〈全タイトル〉と〈序文〉の後に〈論文A〉の全体を繰り返し含むという関係の指定も同様に可能である。

【0043】図1の処理フローの説明に戻ると、処理113で、上のような文書タイプ定義間の対応付け指定がなされ、プログラムがこれを受け取ると、処理は次の処理114に移る。処理114は、上の利用者の指定内容を両文書タイプの識別名のペアによって取り出される形で、外部記憶装置25に保存する。この後、処理は再び処理101に戻って、利用者の要求の指定を待つ。今度は、利用者の要求の指定が文書内容の構造変換であるものとし、処理102から処理121に制御が移ったものとしよう。

【0044】以下、図1の処理121から始まる文書内容の構造変換の処理について説明する。

【0045】まず、処理121では、利用者が変換元／変換先文書例の識別名および変換先文書タイプの識別名を指定するのを受け取る。この後、処理122に進み、指定された変換元文書例134、変換先文書例135および文書タイプ対応情報133を外部記憶装置25から読込む。このとき、指定された変換先文書例がない場合は、新規の文書例とみなして、内容が空の文書例を作成する。指定された変換先文書例135がある場合は、それを読込む。

【0046】文書タイプ対応情報133は、指定された変換元文書例よりわかる変換元文書タイプと指定された変換先文書タイプとのペアによって、外部記憶装置25から読み込むことができる。

【0047】この後、処理123に進み、変換元文書例134の内容を、文書タイプ対応情報133の示す所に従って、構成単位ごとに取り出して、変換先文書例135の対応する構成単位の内容として構造変換し、処理124に進む。

【0048】処理124は、公知の構造文書編集機能であり、ここでは、変換先文書タイプ定義を作成対象の文書の文書タイプ定義とし、処理122で読み込んだ変換先文書例135（または、上述の新規の文書例）を作成中の文書例とし、処理123で作成した変換先文書タイプの構成単位の内容を作成中の文書例の部分の内容として扱うものとする。公知の構造文書編集機能は、このようなインタフェースで、作成対象の文書タイプ定義と作成対象の文書例の構成単位ごとの内容を受け取って、その文書例を作成できる。一構成単位の内容を作成した後、処理125に進み、変換元文書例134の内容がなくなったかどうかを判定する。

【0049】ここで、「終り？」でYESとは、変換元

文書例134の内容がなくなったこと、「終り？」でNOとは、変換元文書例134の内容がまだ残っていることを示す。従って、処理125で、NOの場合には処理123に戻って変換元文書例134のまだ残っている内容の変換処理を続ける。また、処理125で、YESの場合には処理101に戻って利用者の次の要求を受け取る。

【0050】以上により、本発明の文書情報の構造変換方法の処理方法を一とおり説明した。これを本発明の目的と対照してまとめると、次のようである。

【0051】即ち、処理111から処理114において、本発明の第1の目的である、変換元文書タイプ定義および変換先文書タイプ定義の間に利用者が対応付けを指定する方法を提供できることを示した。変換元文書タイプ定義の構成単位と変換先文書タイプ定義の構成単位との間の1:1の対応関係、変換元文書タイプ定義の含む構成単位の一部の構成単位と変換先文書タイプ定義の構成単位の含む構成単位の全体または一部の構成単位との間の1:1の対応関係、変換元文書タイプ定義の含む構成単位の出現順序を変換先文書タイプ定義の構成単位の対応する構成単位の出現順序とが異なるような対応関係などを含めて、種々の対応関係が実現できることも例と共に示した。

【0052】次に、処理121から処理125において、本発明の第2の目的である、利用者が変換元文書例と変換先文書タイプを指定したのに対して、該当する文書タイプ対応情報を参照し、変換先文書例を作成する手段を提供すると共に、作成した変換先文書の構成単位の内容を構造文書編集機能に渡す手段を提供することができ。

【0053】今後、SGML文書の文書タイプ定義のような論理構造の定義とそれに従った文書例の作成、蓄積、交換、利用が普及すること予想される。そこでは、多種類の文書タイプ定義とこの文書タイプ定義ごとにまた多数の文書例が作成される。すなわち、構造上は全く等価で、構造単位の名前が異なる文書タイプ、構成単位の出現順序だけが異なるような文書タイプ、他の文書タイプの構成単位を構成単位として含むような文書タイプ、他の文書タイプの構成単位の繰り返しの構成単位として含むような文書タイプなどがあり得る。そして、これらの文書タイプに従う文書例について、これらの関係を利用した文書例相互間の処理が意味を持つ。即ち、ある学会用に作成した論文を別の学会用の論文に変更すること、ある論文誌の論文のタイトルと著者名と要旨のみを抜き出した要約を作成すること、ある学会の論文を集めて先頭および最後に別形式の文章を付加して1冊の論文誌を構成すること、ある論文誌に記載された複数の論文のタイトルと著者名と要旨のみを抜き出した要約集を作成することなどの要求である。これらの要求は、一括して、文書例の間の文書情報の構造変換の要求であると

いうことができる。

【0054】

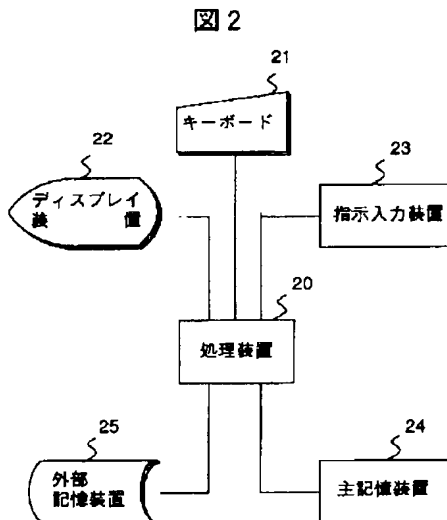
【発明の効果】本発明によれば、変換元文書例の従う文書タイプ定義および変換先文書例の従うべき文書タイプ定義を利用して、変換元文書タイプ定義および変換先文書タイプ定義の間に利用者が対応付けを指定する方法を提供すること、および、実際に与えられた変換元文書例を、このように利用者が指定した対応付けに従って変換し、変換先文書例を作成することによって、この要求を実現することができる。更に、このように変換した文書内容を公知の構造文書編集機能を渡すことによって、変換先の文書を更に、任意に編集したり、印刷したり、割り付けたりすることができる。

【図面の簡単な説明】

【図1】本発明の一実施例における文書情報の構造変換方法を与えるフローチャート。

【図2】本発明の一実施例における文書情報の構造変換

【図2】



方法を与える文書システムのブロック図。

【図3】本発明の一実施例における文書情報の文書タイプ定義の例を与える説明図。

【図4】本発明の一実施例における文書情報の文書タイプ定義の別の例を与える説明図。

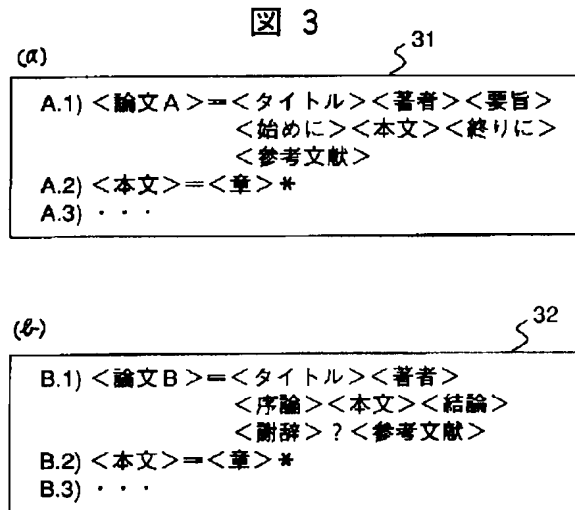
【図5】本発明の一実施例における文書タイプ定義間の利用者による対応付け操作の例を示す説明図。

【図6】本発明の一実施例における文書タイプ定義間の利用者による対応付け操作の別の例を示す説明図。

10 【符号の説明】

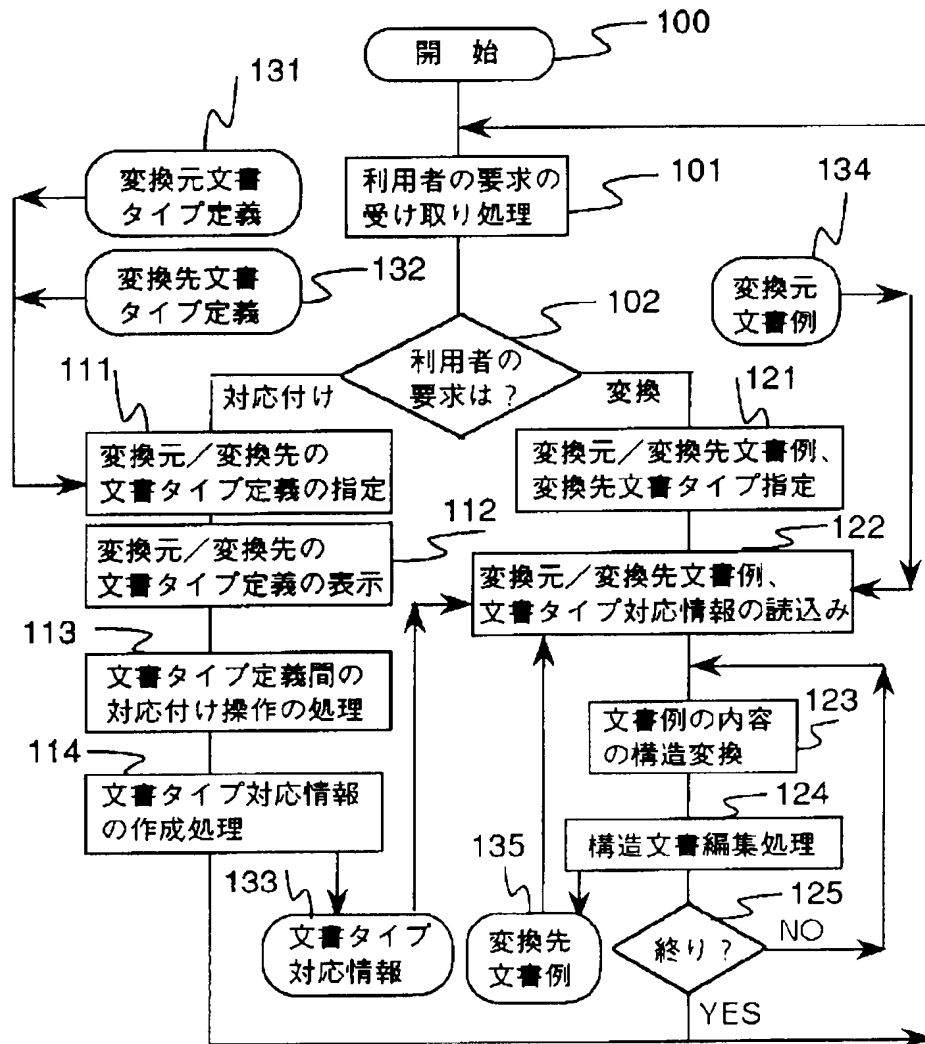
1 1 3…文書タイプ定義間の対応付け操作の処理、1 1 4…文書タイプ対応情報の作成処理、1 2 3…文書例の内容の構造変換、1 2 4…構造文書編集処理、1 3 1…変換元文書タイプ定義、1 3 2…変換先文書タイプ定義、1 3 3…文書タイプ対応情報、1 3 4…変換元文書例、1 3 5…変換先文書例。

【図3】



【図1】

図 1



【图 5】

(a)

(a)

(a)

(a)

(a)

图 5

(a)

(a)

(a)

(a)